

# Multiplicities in Au-Au and Cu-Cu collisions at $\sqrt{s_{NN}} = 62.4$ and 200 GeV

Dariusz Prorok

*Institute of Theoretical Physics, University of Wrocław,  
Pl. Maksa Borna 9, 50-204 Wrocław, Poland*

(Dated: August 2, 2012)

Likelihood ratio tests are performed for the hypothesis that charged-particle multiplicities measured in Au-Au and Cu-Cu collisions at  $\sqrt{s_{NN}} = 62.4$  and 200 GeV are distributed according to the negative binomial form. Results indicate that the hypothesis should be rejected in the all cases of PHENIX-RHIC measurements. Possible explanations of that and of the disagreement with the least-squares fitting method are given.

PACS numbers: 13.85.Hd, 25.75.Ag, 25.75.Gz, 29.85.Fj

## I. INTRODUCTION

The analysis of charged hadron multiplicities in Au-Au and Cu-Cu collisions at  $\sqrt{s_{NN}} = 62.4$  and 200 GeV was done by the PHENIX Collaboration in [1]. It was also claimed there that these multiplicities are distributed according to the negative binomial form. The UA5 Collaboration noticed for the first time that charged-particle multiplicity distributions measured in high energy proton-(anti)proton collisions in limited intervals of pseudo-rapidity have this form [2].

The Negative Binomial Distribution (NBD) is defined as

$$P(n; p, k) = \frac{k(k+1)(k+2)\dots(k+n-1)}{n!} (1-p)^n p^k, \quad (1)$$

where  $n = 0, 1, 2, \dots$ ,  $0 \leq p \leq 1$  and  $k$  is a positive real number. In the application to high energy physics  $n$  has the meaning of the number of charged particles detected in an event. The expected value  $\bar{n}$  and variance  $V(n)$  [10] are expressed as:

$$\bar{n} = \frac{k(1-p)}{p}, \quad V(n) = \frac{k(1-p)}{p^2}. \quad (2)$$

Multiplicity fluctuations are expressed in terms of the scaled variance:

$$\omega = \frac{\langle N_{ch}^2 \rangle - \langle N_{ch} \rangle^2}{\langle N_{ch} \rangle} = \frac{V(n)}{\bar{n}}, \quad (3)$$

where  $N_{ch}$  is the charged particle multiplicity and the last equality is valid only for the whole population (the set of all possible outcomes if the experiment is repeated infinitely many times), assuming that the hypothesis about the NBD is true.

In application to the high energy physics, the parameters  $k, \bar{n}$  instead of  $k, p$  are used usually and

$$\frac{1}{p} = 1 + \frac{\bar{n}}{k} = \omega, \quad (4)$$

which is the scaled variance, Eq. (3). But because the centrality bins have the nonzero width, fluctuations defined by Eq. (3) also include a non-dynamical component. This component is the result of the fluctuations of the geometry of the collisions within a given centrality bin. The geometrical fluctuations were evaluated by the PHENIX Collaboration in [1]. It turned out that those fluctuations can be expressed by a correction factor,  $f_{geo}$ , which is independent of centrality but varies with the collision type. Then the pure scaled variance now representing only dynamical fluctuations, i.e. after subtraction of the geometrical component, can be calculated from the following equation [1]:

$$\omega_{dyn} - 1 = f_{geo} \cdot (\omega - 1). \quad (5)$$

Also parameter  $k$  changes to  $k_{dyn}$  accordingly

$$k_{dyn}^{-1} = f_{geo} \cdot k^{-1} . \quad (6)$$

In this analysis the hypothesis that the charged-particle multiplicities measured in ultra-relativistic heavy-ion collisions are distributed according to the NBD is verified with the use of the maximum likelihood method (ML) and the likelihood ratio test. More details of this approach can be found in Refs. [3–5].

There are two crucial reasons for this approach:

1. The fitted quantity is a probability distribution function (p.d.f.), so the most natural way is to use the ML method, where the likelihood function is constructed directly from the tested p.d.f.. In fact, what is fitted are parameters of the distribution. The fitted values are the *estimators* of these parameters. It is well-known in mathematical statistics that an ML estimator is consistent, asymptotically unbiased and efficient [3, 4, 6]. But even more important is that because of Wilks's theorem (see Appendix B) one can easily define a statistic, the distribution of which converges to a  $\chi^2$  distribution as the number of measurements goes to infinity. Thus for the large sample the goodness-of-fit can be expressed as a  $p$ -value computed with the corresponding  $\chi^2$  distribution.
2. The most commonly used method, the least-squares (LS) method (called also the  $\chi^2$  minimization), has the disadvantage of providing only the qualitative measure of the significance of the fit, in general. Only if observables are represented by Gaussian random variables with known variances, the conclusion about the goodness-of-fit equivalent to that mentioned in the point 1 can be derived [3].

It is worth noting that the ML method with binned data and Poisson fluctuations within a bin was already applied to fitting multiplicity distributions to the NBD but at much lower energies (E-802 Collaboration [7]).

## II. LIKELIHOOD RATIO TEST

The number of charged particles  $N_{ch}$  is assumed to be a random variable with the p.d.f. given by Eq. (1). Each event is treated as an independent observation of  $N_{ch}$  and a set of a given class of events is a sample. For  $N$  events in the class there are  $N$  measurements of  $N_{ch}$ , say  $\mathbf{X} = \{X_1, X_2, \dots, X_N\}$ . Some of these measurements can be equal, *i.e.*  $X_i = X_j$  for  $i \neq j$  can happen. The whole population consists of all possible events with the measurements of 0, 1, 2, ... charged particles and by definition is infinite [11].

Let divide the sample into  $m$  bins characterized by  $Y_i$  - the number of measured charged particles [12] and  $n_i$  - the number of entries in the  $i$ th bin,  $N = \sum_{i=1}^m n_i$  (details of the theoretical framework of this Section can be found in Refs. [3–5]). Then the expectation value of the number of events in the  $i$ th bin can be written as

$$\nu_i(\nu_{tot}, p, k) = \nu_{tot} \cdot P(Y_i; p, k) , \quad (7)$$

where  $\nu_{tot}$  is the expected number of all events in the sample,  $\nu_{tot} = \sum_{i=1}^m \nu_i$ . This is because one can treat the number of events in the sample  $N$  also as a random variable with its own distribution - Poisson one. Generally, the whole histogram can be treated as one measurement of  $m$ -dimensional random vector  $\mathbf{n} = (n_1, \dots, n_m)$  which has a multinomial distribution, so the joint p.d.f. for the measurement of  $N$  and  $\mathbf{n}$  can be converted to the form [3, 5]:

$$f(\mathbf{n}; \nu_1, \dots, \nu_m) = \prod_{i=1}^m \frac{\nu_i^{n_i}}{n_i!} \exp(-\nu_i) . \quad (8)$$

Since now  $f(\mathbf{n}; \nu_1, \dots, \nu_m)$  is the p.d.f. for one measurement,  $f$  is also the likelihood function

$$L(\mathbf{n} \mid \nu_1, \dots, \nu_m) = f(\mathbf{n}; \nu_1, \dots, \nu_m) . \quad (9)$$

With the use of Eq. (7) the corresponding likelihood function can be written as

$$L(\mathbf{n} \mid \nu_{tot}, p, k) = L(\mathbf{n} \mid \nu_1(\nu_{tot}, p, k), \dots, \nu_m(\nu_{tot}, p, k)) . \quad (10)$$

Then the likelihood ratio is defined as

$$\lambda = \frac{L(\mathbf{n} | \hat{\nu}_{tot}, \hat{p}, \hat{k})}{L(\mathbf{n} | \check{\nu}_1, \dots, \check{\nu}_m)} = \frac{L(\mathbf{n} | \hat{\nu}_{tot}, \hat{p}, \hat{k})}{L(\mathbf{n} | n_1, \dots, n_m)}. \quad (11)$$

where  $\hat{\nu}_{tot}$ ,  $\hat{p}$  and  $\hat{k}$  are the ML estimates of  $\nu_{tot}$ ,  $p$  and  $k$  with the likelihood function given by Eq. (10) and  $\check{\nu}_i = n_i$ ,  $i = 1, 2, \dots, m$  are the ML estimates of  $\nu_i$  treated as free parameters. Note that since the denominator in Eq. (11) does not depend on parameters, the log-ratio defined as

$$\begin{aligned} \ln \lambda(\nu_{tot}, p, k) &= \ln \frac{L(\mathbf{n} | \nu_{tot}, p, k)}{L(\mathbf{n} | n_1, \dots, n_m)} \\ &= - \sum_{i=1}^m \left( n_i \ln \frac{n_i}{\nu_i} + \nu_i - n_i \right) \\ &= -\nu_{tot} + N - \sum_{i=1}^m n_i \ln \frac{n_i}{\nu_i}, \end{aligned} \quad (12)$$

where  $\nu_i$  are expressed by Eq. (7), can be used to find the ML estimates of  $\nu_{tot}$ ,  $p$  and  $k$ . The values  $\hat{\nu}_{tot}$ ,  $\hat{p}$  and  $\hat{k}$  for which  $\lambda(\nu_{tot}, p, k)$  has its maximum are the maximum likelihood estimates of parameters  $\nu_{tot}$ ,  $p$  and  $k$ . Further, the statistic given by

$$\chi^2 = -2 \ln \lambda(\hat{\nu}_{tot}, \hat{p}, \hat{k}) = 2 \sum_{i=1}^m \left( n_i \ln \frac{n_i}{\hat{\nu}_i} + \hat{\nu}_i - n_i \right) \quad (13)$$

approaches the  $\chi^2$  distribution asymptotically, *i.e.* as the number of measurements, here the number of events  $N$ , goes to infinity (the consequence of the Wilks's theorem, see Appendix B). The values  $\hat{\nu}_i$  are the estimates of  $\nu_i$  given by

$$\hat{\nu}_i = \hat{\nu}_{tot} \cdot P(Y_i; \hat{p}, \hat{k}) \quad (14)$$

and if one assumes that  $\nu_{tot}$  does not depend on  $p$  and  $k$  then  $\hat{\nu}_{tot} = N$ . For such a case

$$\sum_{i=1}^m \hat{\nu}_i = \sum_{i=1}^m n_i \quad (15)$$

and Eq. (13) becomes

$$\chi^2(\hat{p}, \hat{k}) = -2 \ln \lambda = 2 \sum_{i=1}^m n_i \ln \frac{n_i}{\hat{\nu}_i}. \quad (16)$$

Also then one can just put  $\nu_{tot} = N$  and Eq. (12) can be rewritten as

$$\begin{aligned} \ln \lambda(p, k) &= N \cdot \ln N - \sum_{i=1}^m n_i \ln n_i + \sum_{i=1}^m n_i \ln P(Y_i; p, k) \\ &= - \sum_{i=1}^m n_i \ln \frac{n_i}{N} + N \sum_{i=1}^m \frac{n_i}{N} \ln P(Y_i; p, k) \\ &= -N \sum_{i=1}^m P_i^{ex} \ln P_i^{ex} + N \sum_{i=1}^m P_i^{ex} \ln P(Y_i; p, k), \end{aligned} \quad (17)$$

where  $P_i^{ex} = n_i/N$ . Note that the maximum of  $\ln \lambda$  is the minimum of  $\chi^2 = -2 \ln \lambda$ , so from Eqs. (16) and (17) one arrives at

$$\chi_{min}^2 = -2 N \sum_{i=1}^m P_i^{ex} \ln \frac{P(Y_i; \hat{p}, \hat{k})}{P_i^{ex}}. \quad (18)$$

It can be proven that one of the necessary conditions for the existence of the maximum is (see Appendix A for details):

$$\bar{n} = \langle N_{ch} \rangle, \quad (19)$$

*i.e.* the distribution average has to be equal to the experimental average. This is very good because  $\langle N_{ch} \rangle$  is what is called in statistics *a sample mean*. The sample mean is an estimator for the expectation value of the random variable, which is consistent and unbiased [3]. In other words the ML estimator of  $\bar{n}$  is  $\langle N_{ch} \rangle$  ( $\hat{n} = \langle N_{ch} \rangle$ ).

### III. RESULTS AND DISCUSSION

The method described in Section II requires that all bins in a given data set have the width equal to 1, so as the experimental probability  $P_i^{ex}$  to measure a signal in the  $i$ th bin was equivalent to the probability of the measurement of  $(i - 1)$  charged particles (the first bin is the bin of 0 charged particles detected). This is fulfilled for all bins of the considered data sets

Since the test statistic  $-2 \ln \lambda$  has a  $\chi^2$  distribution approximately in the large sample limit, it can be used as a test of the goodness-of-fit. The result of the test is given by the so-called  $p$ -value which is the probability of obtaining the value of the statistic, Eq. (13), equal to or greater then the value just obtained for the present data set, when repeating the whole experiment many times:

$$p = P(\chi^2 \geq \chi_{min}^2; n_d) = \int_{\chi_{min}^2}^{\infty} f(z; n_d) dz, \quad (20)$$

where  $f(z; n_d)$  is the  $\chi^2$  p.d.f. and  $n_d$  the number of degrees of freedom,  $n_d = m - 2$  here.

The results of the analysis are presented in Tables I-VIII and illustrated with Figs. 1-6. In fact the whole analysis was done for the two kinds of histograms: (i) bins with the number of entries  $n_i \leq 5$  excluded, Tables I, III, V and VII; (ii) bins with the number of entries  $n_i \leq 60$  excluded, Tables II, IV, VI and VIII. In practice this corresponds to cutting off less (i) or more (ii) the tails of the full measured histograms. The tails breaks the visual agreement between the data and the NBD, cf. Figs. 1 and 2. The condition that only bins with  $n_i > 5$  are taken into account is the minimal condition imposed on a histogram to do any statistical inference without Monte Carlo simulations [3]. The condition (ii) corresponds to the choice made originally by the PHENIX Collaboration in their analysis [1]. It has turned out that the results of this analysis are qualitatively the same for both choices.

As one can see, the hypothesis in question should be rejected in all considered cases. But it was claimed that charged-particle multiplicities measured in Au-Au and Cu-Cu collisions at  $\sqrt{s_{NN}} = 62.4$  and 200 GeV are distributed according to the NBD [1]. However that conclusion was the result of the application of the LS method. Therefore it seems to be reasonable to check what are the values of the LS test statistic at the ML estimators listed in the third and fourth columns of Tables I-VIII. For the sample described in Sect. II one can define the LS test statistic (commonly called the  $\chi^2$  function) as:

$$\chi_{LS}^2(\bar{n}, k) = \sum_{i=1}^m \frac{(P_i^{ex} - P(Y_i; \bar{n}, k))^2}{err_i^2}, \quad (21)$$

where  $err_i$  is the uncertainty of the  $i$ th measurement. Here this function **is not minimized** with respect to  $\bar{n}$  and  $k$  (or  $p$  and  $k$ ) as in the LS method but is calculated at ML estimates of  $\bar{n}$  and  $k$ . This is allowed in statistics and is equivalent to test a single point in the parameter space. Then the tested point might not be the best estimate of the true value but the hypothesis in question becomes the hypothesis only about a particular distribution (a *simple* hypothesis). One can see from the ninth column of Tables I-VIII that  $\chi_{LS}^2/n_d$  values are significant for almost all centrality classes, what agrees with the results of Ref. [1]. But this contradicts the results of the likelihood ratio test, which are expressed by  $\chi^2/n_d$  and  $p$ -values listed in the seventh and eight columns of Tables I-VIII.

TABLE I: Results of fitting multiplicity distributions measured by the PHENIX Collaboration in Au-Au collisions at  $\sqrt{s_{NN}} = 200$  GeV,  $f_{geo} = 0.37 \pm 0.027$  [1]. Fitting ranges are limited to the bins with  $n_i > 5$ , where  $n_i$  is the number of events in the  $i$ th bin.

Centrality [%]	N	$\hat{k}$	$\hat{n}$	$1/\hat{k}_{dyn}$	$\omega_{dyn}$	$\chi^2/n_d$	p-value [%]	$\chi^2_{LS}/n_d$ with errors:	
						$\chi^2$ ( $n_d$ )		quadrature sum	statistical only
0-5	653145	270.0 $\pm 2.5$	61.85 $\pm 0.01$	$1.37 \cdot 10^{-3}$ $\pm 0.10 \cdot 10^{-3}$	1.08 $\pm 0.01$	23.73 1756.0 (74)	0	0.98	9.8
5-10	657944	163.4 $\pm 1.2$	53.91 $\pm 0.01$	$2.26 \cdot 10^{-3}$ $\pm 0.17 \cdot 10^{-3}$	1.12 $\pm 0.01$	9.12 592.7 (65)	0	0.69	6.9
10-15	658739	112.5 $\pm 0.7$	46.50 $\pm 0.01$	$3.29 \cdot 10^{-3}$ $\pm 0.24 \cdot 10^{-3}$	1.15 $\pm 0.01$	11.5 795.5 (69)	0	0.66	6.6
15-20	659607	85.1 $\pm 0.5$	39.72 $\pm 0.01$	$4.35 \cdot 10^{-3}$ $\pm 0.32 \cdot 10^{-3}$	1.17 $\pm 0.01$	8.9 585.8 (66)	0	0.52	5.2
20-25	658785	67.6 $\pm 0.4$	33.56 $\pm 0.01$	$5.48 \cdot 10^{-3}$ $\pm 0.40 \cdot 10^{-3}$	1.18 $\pm 0.01$	13.5 848.8 (63)	0	0.46	4.6
25-30	659632	56.7 $\pm 0.3$	28.01 $\pm 0.01$	$6.52 \cdot 10^{-3}$ $\pm 0.48 \cdot 10^{-3}$	1.18 $\pm 0.01$	10.9 640.6 (59)	0	0.37	3.7
30-35	659303	47.4 $\pm 0.3$	23.02 $\pm 0.01$	$7.81 \cdot 10^{-3}$ $\pm 0.57 \cdot 10^{-3}$	1.18 $\pm 0.01$	7.9 429.9 (54)	0	0.31	3.1
35-40	661174	40.5 $\pm 0.2$	18.64 $\pm 0.01$	$9.13 \cdot 10^{-3}$ $\pm 0.67 \cdot 10^{-3}$	1.17 $\pm 0.01$	8.5 389.7 (46)	0	0.37	3.7
40-45	661599	34.0 $\pm 0.2$	14.84 $\pm 0.01$	$1.09 \cdot 10^{-2}$ $\pm 0.80 \cdot 10^{-3}$	1.16 $\pm 0.01$	7.3 301.0 (41)	0	0.35	3.5
45-50	661765	27.3 $\pm 0.2$	11.57 $\pm 0.005$	$1.35 \cdot 10^{-2}$ $\pm 0.99 \cdot 10^{-3}$	1.16 $\pm 0.01$	10.5 390.2 (37)	0	0.92	9.2
50-55	662114	21.3 $\pm 0.1$	8.82 $\pm 0.004$	$1.74 \cdot 10^{-2}$ $\pm 0.13 \cdot 10^{-2}$	1.15 $\pm 0.01$	38.8 1436.4 (37)	0	12.06	120.6

TABLE II: Results of fitting multiplicity distributions measured by the PHENIX Collaboration in Au-Au collisions at  $\sqrt{s_{NN}} = 200$  GeV,  $f_{geo} = 0.37 \pm 0.027$  [1]. Fitting ranges are limited to the bins with  $n_i > 60$ , where  $n_i$  is the number of events in the  $i$ th bin.

Centrality [%]	N	$\hat{k}$	$\hat{n}$	$1/\hat{k}_{dyn}$	$\omega_{dyn}$	$\chi^2/n_d$	p-value [%]	$\chi^2_{LS}/n_d$ with errors:	
						$\chi^2$ ( $n_d$ )		quadrature sum	statistical only
0-5	652579	289.0 $\pm 2.9$	61.86 $\pm 0.01$	$1.28 \cdot 10^{-3}$ $\pm 0.94 \cdot 10^{-4}$	1.08 $\pm 0.01$	20.0 1160.2 (58)	0	0.57	5.7
5-10	657571	168.1 $\pm 1.2$	53.91 $\pm 0.01$	$2.20 \cdot 10^{-3}$ $\pm 0.16 \cdot 10^{-3}$	1.12 $\pm 0.01$	20.56 1151.6 (56)	0	0.61	6.1
10-15	658258	116.4 $\pm 0.7$	46.50 $\pm 0.01$	$3.18 \cdot 10^{-3}$ $\pm 0.23 \cdot 10^{-3}$	1.15 $\pm 0.01$	18.4 991.7 (54)	0	0.53	5.3
15-20	659302	86.9 $\pm 0.5$	39.72 $\pm 0.01$	$4.26 \cdot 10^{-3}$ $\pm 0.31 \cdot 10^{-3}$	1.17 $\pm 0.01$	12.6 667.5 (53)	0	0.43	4.3
20-25	658461	69.1 $\pm 0.4$	33.56 $\pm 0.01$	$5.36 \cdot 10^{-3}$ $\pm 0.39 \cdot 10^{-3}$	1.18 $\pm 0.01$	12.3 604.7 (49)	0	0.34	3.4
25-30	659337	57.9 $\pm 0.3$	28.0 $\pm 0.01$	$6.39 \cdot 10^{-3}$ $\pm 0.47 \cdot 10^{-3}$	1.18 $\pm 0.01$	10.4 469.1 (45)	0	0.28	2.8
30-35	659021	48.3 $\pm 0.3$	23.02 $\pm 0.01$	$7.66 \cdot 10^{-3}$ $\pm 0.56 \cdot 10^{-3}$	1.18 $\pm 0.01$	8.6 351.02 (41)	0	0.16	1.6
35-40	660937	41.3 $\pm 0.2$	18.64 $\pm 0.01$	$8.96 \cdot 10^{-3}$ $\pm 0.66 \cdot 10^{-3}$	1.17 $\pm 0.01$	7.6 280.3 (37)	0	0.19	1.9
40-45	661422	34.6 $\pm 0.2$	14.84 $\pm 0.01$	$1.07 \cdot 10^{-2}$ $\pm 0.78 \cdot 10^{-3}$	1.16 $\pm 0.01$	7.9 260.3 (33)	0	0.21	2.1
45-50	661577	27.9 $\pm 0.2$	11.56 $\pm 0.005$	$1.33 \cdot 10^{-2}$ $\pm 0.97 \cdot 10^{-3}$	1.15 $\pm 0.01$	10.0 279.9 (28)	0	0.23	2.3
50-55	661877	21.9 $\pm 0.1$	8.81 $\pm 0.004$	$1.69 \cdot 10^{-2}$ $\pm 0.12 \cdot 10^{-2}$	1.15 $\pm 0.01$	40.0 959.2 (24)	0	0.30	3.0

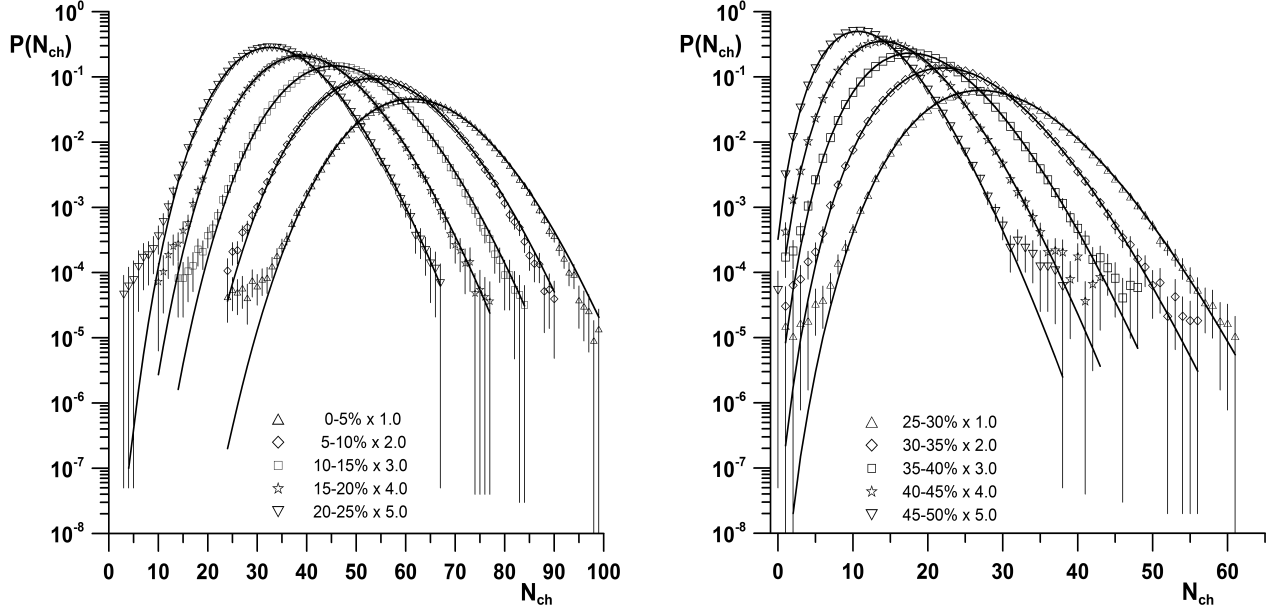


FIG. 1: Uncorrected multiplicity distributions of charged hadrons for 200 GeV Au-Au collisions [1] within ranges limited to the bins with  $n_i > 5$ . The lines are fits to the NBD. The data are scaled by the amounts in the legend. Errors represent the statistical and systematic errors added in quadrature.

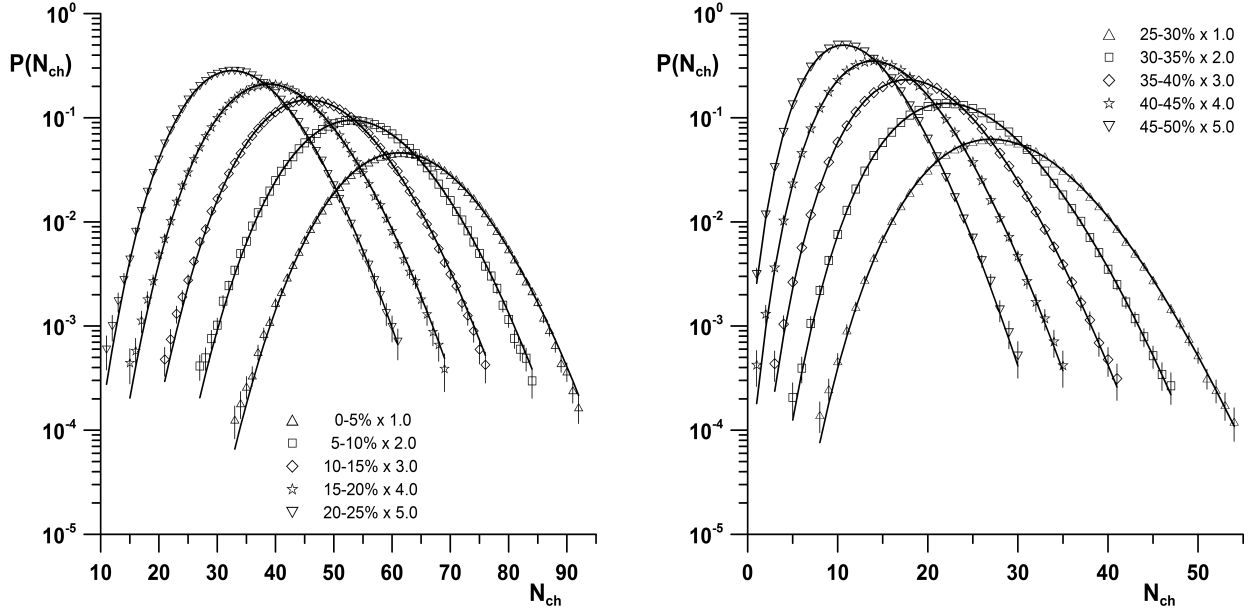


FIG. 2: Uncorrected multiplicity distributions of charged hadrons for 200 GeV Au-Au collisions [1] within ranges limited to the bins with  $n_i > 60$ . The lines are fits to the NBD. The data are scaled by the amounts in the legend. Errors represent the statistical and systematic errors added in quadrature.

However, the ML method and the likelihood ratio test have much more solid mathematical foundations as it has been explained in Sect. I. Thus the hypothesis about the NBD should be rejected on the basis of obtained values of  $\chi^2/n_d$  and  $p$ -values.

The crucial question is now why the conclusions from  $\chi^2$  and  $\chi^2_{LS}$  test statistics are entirely opposite for PHENIX measurements? The main difference between both statistics is that  $\chi^2$  depends explicitly on the number of events but  $\chi^2_{LS}$  does not. On opposite,  $\chi^2$  does not depend on the actual errors but  $\chi^2_{LS}$  does. In fact,  $\chi^2$  statistic implicitly

TABLE III: Results of fitting multiplicity distributions measured by the PHENIX Collaboration in Au-Au collisions at  $\sqrt{s_{NN}} = 62.4$  GeV,  $f_{geo} = 0.33 \pm 0.031$  [1]. Fitting ranges are limited to the bins with  $n_i > 5$ , where  $n_i$  is the number of events in the  $i$ th bin.

Centrality [%]	N	$\hat{k}$	$\hat{n}$	$1/\hat{k}_{dyn}$	$\omega_{dyn}$	$\chi^2/n_d$	p-value [%]	$\chi^2_{LS}/n_d$ with errors:	
						$\chi^2$ ( $n_d$ )		quadrature sum	statistical only
0-5	607155	225.2 $\pm 2.5$	44.67 $\pm 0.01$	$1.47 \cdot 10^{-3}$ $\pm 0.14 \cdot 10^{-3}$	1.07 $\pm 0.01$	2.37 139.6 (59)	$1.7 \cdot 10^{-8}$	0.18	1.8
5-10	752392	142.3 $\pm 1.1$	37.96 $\pm 0.01$	$2.32 \cdot 10^{-3}$ $\pm 0.22 \cdot 10^{-3}$	1.09 $\pm 0.01$	2.44 131.9 (54)	$1.9 \cdot 10^{-8}$	0.11	1.1
10-15	752837	115.2 $\pm 0.9$	31.53 $\pm 0.01$	$2.87 \cdot 10^{-3}$ $\pm 0.27 \cdot 10^{-3}$	1.09 $\pm 0.01$	2.06 107.1 (52)	$1.1 \cdot 10^{-5}$	0.13	1.3
15-20	752553	88.0 $\pm 0.6$	26.07 $\pm 0.01$	$3.75 \cdot 10^{-3}$ $\pm 0.35 \cdot 10^{-3}$	1.10 $\pm 0.01$	1.86 87.3 (47)	$3.2 \cdot 10^{-4}$	0.13	1.3
20-25	752296	68.5 $\pm 0.5$	21.35 $\pm 0.01$	$4.82 \cdot 10^{-3}$ $\pm 0.45 \cdot 10^{-3}$	1.10 $\pm 0.01$	2.63 113.2 (43)	$3.1 \cdot 10^{-8}$	0.21	2.1
25-30	752183	53.2 $\pm 0.4$	17.30 $\pm 0.01$	$6.21 \cdot 10^{-3}$ $\pm 0.59 \cdot 10^{-3}$	1.11 $\pm 0.01$	2.75 107.3 (39)	$2.7 \cdot 10^{-8}$	0.23	2.3
30-35	751375	40.1 $\pm 0.3$	13.84 $\pm 0.005$	$8.22 \cdot 10^{-3}$ $\pm 0.77 \cdot 10^{-3}$	1.11 $\pm 0.01$	2.97 103.9 (35)	$9.6 \cdot 10^{-9}$	0.25	2.5
35-40	751661	31.7 $\pm 0.2$	10.89 $\pm 0.004$	$1.04 \cdot 10^{-2}$ $\pm 0.98 \cdot 10^{-3}$	1.11 $\pm 0.01$	6.72 194.9 (29)	0	0.16	1.6
40-45	750884	25.1 $\pm 0.2$	8.42 $\pm 0.004$	$1.31 \cdot 10^{-2}$ $\pm 0.12 \cdot 10^{-2}$	1.11 $\pm 0.01$	37.5 937.4 (25)	0	40.36	403.6
45-50	751421	21.8 $\pm 0.2$	6.41 $\pm 0.003$	$1.51 \cdot 10^{-2}$ $\pm 0.14 \cdot 10^{-2}$	1.10 $\pm 0.01$	209.0 4806.8 (23)	0	285.9	2859.5

assumes errors of the type  $\sqrt{n_i}$ , what is the straightforward result of the form of the likelihood function, Eqs. (8) and (9), namely the product of Poisson distributions. In the PHENIX analysis [1] errors  $err_i$  in Eq. (21) are represented by the quadrature sum of the statistical and systematic components. And the systematic errors were estimated as 3.0 times the statistical errors [8]. This causes that  $err_i^2 = 10 \cdot \sigma_{stat,i}^2$ , where  $\sigma_{stat,i}$  is the statistical error of the  $i$ th



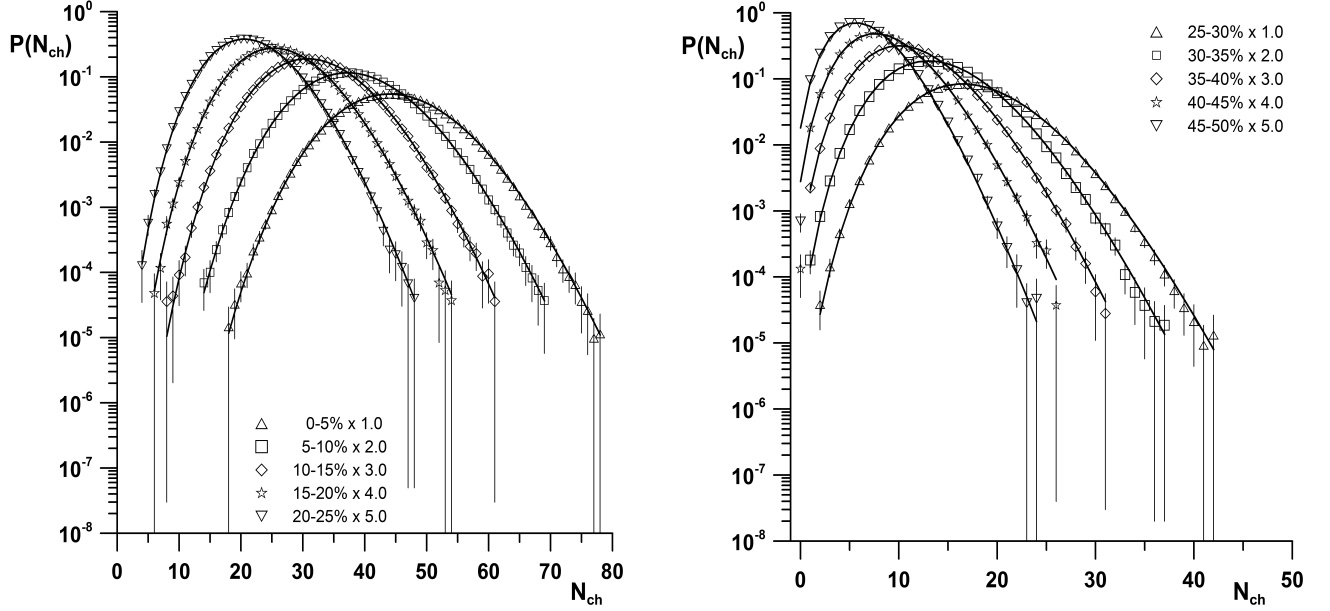


FIG. 3: Uncorrected multiplicity distributions of charged hadrons for 62.4 GeV Au-Au collisions [1] within ranges limited to the bins with  $n_i > 5$ . The lines are fits to the NBD. The data are scaled by the amounts in the legend. Errors represent the statistical and systematic errors added in quadrature.

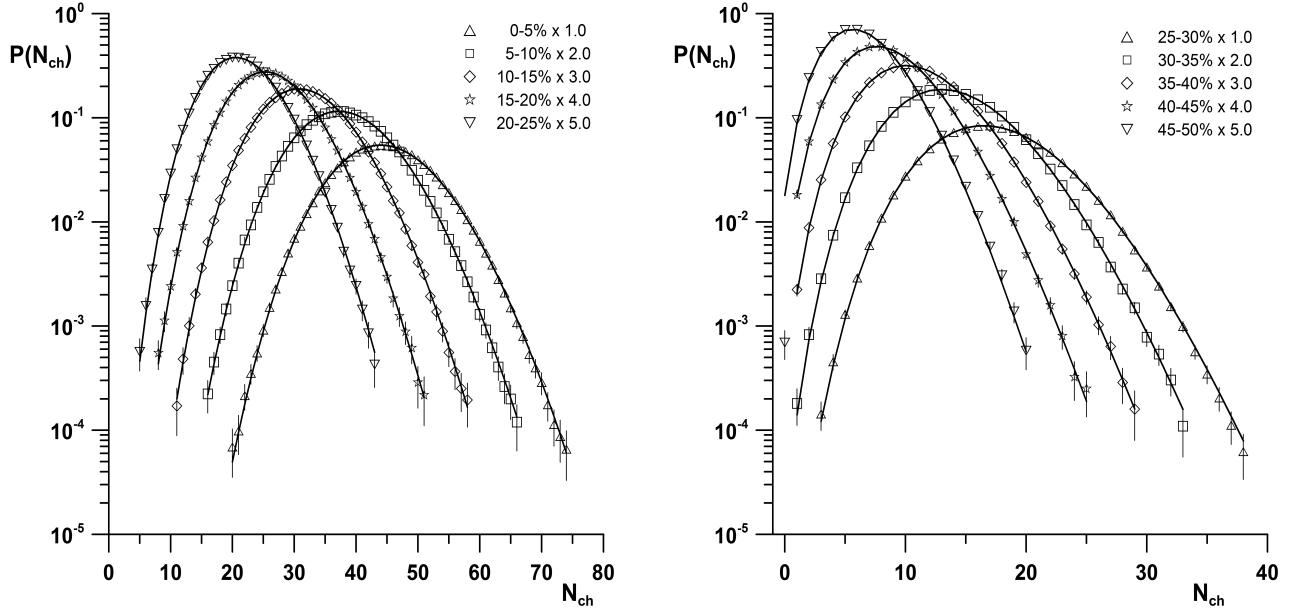


FIG. 4: Uncorrected multiplicity distributions of charged hadrons for 62.4 GeV Au-Au collisions [1] within ranges limited to the bins with  $n_i > 60$ . The lines are fits to the NBD. The data are scaled by the amounts in the legend. Errors represent the statistical and systematic errors added in quadrature.

measurement. Hence the values of  $\chi^2_{LS}/n_d$  in last columns of Tables I-VIII are 10 times greater than in preceding columns.

In principle, the accuracy with which experimental distributions approximate the NBD should increase with the sample size because if the hypothesis is true the postulated form of distribution is exact for the whole population. So with the growing number of events, the experimental distribution should be closer to the postulated one. This is also seen in the form of  $\chi^2_{min}$ , Eq. (18), where the linear dependence on  $N$  is explicit. To keep  $\chi^2_{min}$  at least constant

TABLE IV: Results of fitting multiplicity distributions measured by the PHENIX Collaboration in Au-Au collisions at  $\sqrt{s_{NN}} = 62.4$  GeV,  $f_{geo} = 0.33 \pm 0.031$  [1]. Fitting ranges are limited to the bins with  $n_i > 60$ , where  $n_i$  is the number of events in the  $i$ th bin.

Centrality [%]	N	$\hat{k}$	$\hat{n}$	$1/\hat{k}_{dyn}$	$\omega_{dyn}$	$\chi^2/n_d$	p-value [%]	$\chi^2_{LS}/n_d$ with errors:	
						$\chi^2$ ( $n_d$ )		quadrature sum	statistical only
0-5	607075	227.9 $\pm 2.5$	44.67 $\pm 0.01$	$1.45 \cdot 10^{-3}$ $\pm 0.14 \cdot 10^{-3}$	1.06 $\pm 0.01$	5.55 294.3 (53)	0	0.19	1.9
5-10	752263	143.9 $\pm 1.1$	37.96 $\pm 0.01$	$2.29 \cdot 10^{-3}$ $\pm 0.22 \cdot 10^{-3}$	1.09 $\pm 0.01$	7.80 382.4 (49)	0	0.12	1.2
10-15	752739	116.2 $\pm 0.9$	31.53 $\pm 0.01$	$2.84 \cdot 10^{-3}$ $\pm 0.27 \cdot 10^{-3}$	1.09 $\pm 0.01$	5.67 260.8 (46)	0	0.13	1.3
15-20	752492	88.5 $\pm 0.6$	26.07 $\pm 0.01$	$3.73 \cdot 10^{-3}$ $\pm 0.35 \cdot 10^{-3}$	1.10 $\pm 0.01$	5.97 250.9 (42)	0	0.11	1.1
20-25	752182	69.2 $\pm 0.5$	21.35 $\pm 0.01$	$4.77 \cdot 10^{-3}$ $\pm 0.45 \cdot 10^{-3}$	1.10 $\pm 0.01$	10.2 377.2 (37)	0	0.22	2.2
25-30	752095	53.6 $\pm 0.4$	17.30 $\pm 0.01$	$6.16 \cdot 10^{-3}$ $\pm 0.58 \cdot 10^{-3}$	1.11 $\pm 0.01$	8.2 279.2 (34)	0	0.23	2.3
30-35	751324	40.3 $\pm 0.3$	13.84 $\pm 0.005$	$8.19 \cdot 10^{-3}$ $\pm 0.77 \cdot 10^{-3}$	1.11 $\pm 0.01$	7.40 229.3 (31)	0	0.26	2.6
35-40	751639	31.8 $\pm 0.2$	10.89 $\pm 0.004$	$1.04 \cdot 10^{-2}$ $\pm 0.98 \cdot 10^{-3}$	1.11 $\pm 0.01$	9.43 254.7 (27)	0	0.15	1.5
40-45	750852	25.2 $\pm 0.2$	8.42 $\pm 0.004$	$1.31 \cdot 10^{-2}$ $\pm 0.12 \cdot 10^{-2}$	1.11 $\pm 0.01$	50.7 1166.3 (23)	0	0.22	2.2
45-50	751348	22.0 $\pm 0.2$	6.41 $\pm 0.003$	$1.50 \cdot 10^{-2}$ $\pm 0.14 \cdot 10^{-2}$	1.10 $\pm 0.01$	259.8 4936.4 (19)	0	343.1	3431.1

when  $N$  (the sample size) is growing the relative differences between  $P(Y_i)$  and  $P_i^{ex}$  have to decrease. The results of the likelihood ratio test have revealed that the discrepancies between the measured and postulated distributions are too big for these sample sizes to accept the NBD hypothesis.

Another surprising point is the application of the LS test statistic with the ML estimators of the parameters to the

TABLE V: Results of fitting multiplicity distributions measured by the PHENIX Collaboration in Cu-Cu collisions at  $\sqrt{s_{NN}} = 200$  GeV,  $f_{geo} = 0.40 \pm 0.047$  [1]. Fitting ranges are limited to the bins with  $n_i > 5$ , where  $n_i$  is the number of events in the  $i$ th bin.

Centrality [%]	N	$\hat{k}$	$\hat{n}$	$1/\hat{k}_{dyn}$	$\omega_{dyn}$	$\chi^2/n_d$	p-value [%]	$\chi^2_{LS}/n_d$ with errors:	
						$\chi^2$ ( $n_d$ )		quadrature sum	statistical only
0-5	368510	59.6	19.80	$6.72 \cdot 10^{-3}$	1.13	94.8	0	2.1	21.2
		$\pm 0.6$	$\pm 0.01$	$\pm 0.79 \cdot 10^{-3}$	$\pm 0.02$	3887.0 (41)			
5-10	369206	49.6	16.74	$8.06 \cdot 10^{-3}$	1.13	16.5	0	0.66	6.6
		$\pm 0.5$	$\pm 0.01$	$\pm 0.95 \cdot 10^{-3}$	$\pm 0.02$	628.5 (38)			
10-15	369945	41.5	14.05	$9.64 \cdot 10^{-3}$	1.14	6.8	0	0.38	3.8
		$\pm 0.4$	$\pm 0.01$	$\pm 0.11 \cdot 10^{-2}$	$\pm 0.02$	225.5 (33)			
15-20	370066	34.5	11.78	$1.16 \cdot 10^{-2}$	1.14	3.0	$5.8 \cdot 10^{-8}$	0.24	2.4
		$\pm 0.3$	$\pm 0.01$	$\pm 0.14 \cdot 10^{-2}$	$\pm 0.02$	92.0 (31)			
20-25	371877	29.2	9.81	$1.37 \cdot 10^{-2}$	1.13	6.6	0	3.4	33.6
		$\pm 0.3$	$\pm 0.01$	$\pm 0.16 \cdot 10^{-2}$	$\pm 0.02$	186.0 (28)			
25-30	368876	24.9	8.14	$1.60 \cdot 10^{-2}$	1.13	19.3	0	11.5	114.9
		$\pm 0.2$	$\pm 0.01$	$\pm 0.19 \cdot 10^{-2}$	$\pm 0.02$	502.4 (26)			
30-35	368072	21.9	6.72	$1.83 \cdot 10^{-2}$	1.12	65.6	0	42.3	422.5
		$\pm 0.2$	$\pm 0.005$	$\pm 0.22 \cdot 10^{-2}$	$\pm 0.01$	1704.8 (26)			

data samples. The results are presented in the last two columns of Tables I-VIII. For the choice (ii) which corresponds to PHENIX analysis [1], Tables II, IV, VI and VIII, the  $\chi^2_{LS}/n_d$  values obtained here are lower than corresponding ones in Ref. [1]. Values of the parameters  $\hat{k}, \hat{n}$  are also different from those in Ref. [1], what has resulted in slightly different (1 – 3% lower) values of the scaled variance  $\omega_{dyn}$ , see Figs. 7 and 8. To make the comparison easier also values of  $\hat{k}_{dyn}^{-1}$  are presented in the fifth column of Tables I-VIII. Generally,  $\hat{n}$  is greater but the difference does not exceed 10% and decreases with the centrality.  $\hat{k}_{dyn}^{-1}$  is smaller, especially for case (ii) and the difference also decreases with the centrality; from about 20 – 30% for the least central classes to about 5 – 10% for the most central ones.

#### IV. CONCLUSIONS

The main conclusion is that the hypothesis of the NBD of charged-particle multiplicities measured by the PHENIX Collaboration in Au-Au and Cu-Cu collisions at  $\sqrt{s_{NN}} = 62.4$  and 200 GeV should be rejected for all centrality classes. This is the result of likelihood ratio tests performed for the corresponding data samples. The significant systematic errors are the reasons for acceptable values of the least squares test statistic, which is easily seen from the last two columns of Tables I-VIII. If the systematic errors are not taken into account, then  $\chi^2_{LS}/n_d$  values are greater

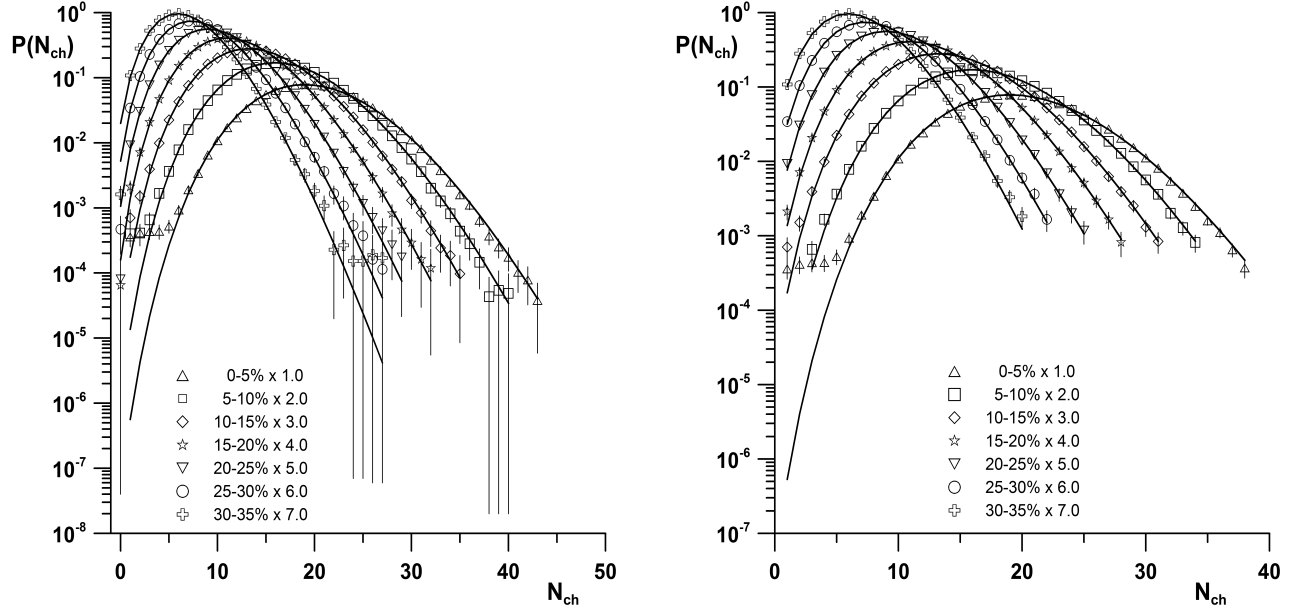


FIG. 5: Uncorrected multiplicity distributions of charged hadrons for 200 GeV Cu-Cu collisions [1] within ranges limited to the bins with  $n_i > 5$  (left) and  $n_i > 60$  (right). The lines are fits to the NBD. The data are scaled by the amounts in the legend. Errors represent the statistical and systematic errors added in quadrature.

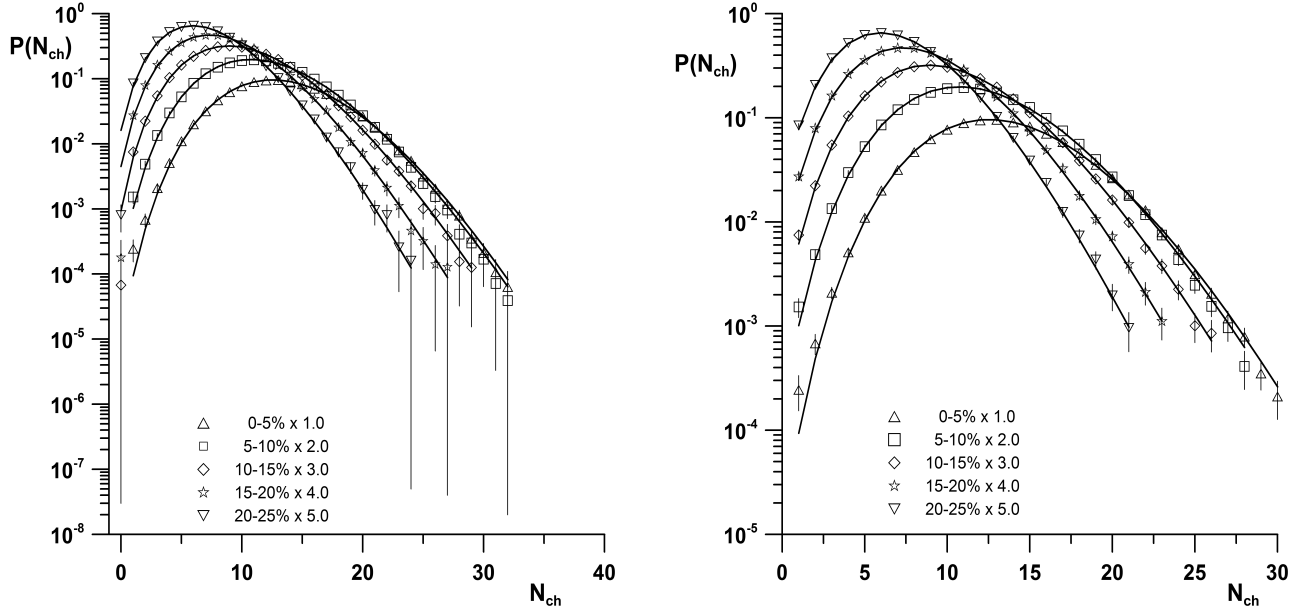


FIG. 6: Uncorrected multiplicity distributions of charged hadrons for 62.4 GeV Cu-Cu collisions [1] within ranges limited to the bins with  $n_i > 5$  (left) and  $n_i > 60$  (right). The lines are fits to the NBD. The data are scaled by the amounts in the legend. Errors represent the statistical and systematic errors added in quadrature.

than 1 in all considered cases and substantially greater than 1 in most of them.

The second, more general conclusion is that the scale of errors should be somehow related to the sample size to make the statistical inference trustworthy in the case of the LS method. This is because too big or too small errors cause the false inference in this case. Hence errors of the type  $\sqrt{n_i}$  could be "a frame of reference". Note that both types of errors are such in all considered PHENIX cases; for  $n_i$  statistical errors are equal to  $\sqrt{n_i}$  exactly and systematic errors are 3.0 times the statistical ones [1, 8], but the last factor causes that the total error is  $\sqrt{10} \cdot \sigma_{stat,i} \approx 3.0 \cdot \sigma_{stat,i} = 3.0 \cdot \sqrt{n_i}$ .

TABLE VI: Results of fitting multiplicity distributions measured by the PHENIX Collaboration in Cu-Cu collisions at  $\sqrt{s_{NN}} = 200$  GeV,  $f_{geo} = 0.40 \pm 0.047$  [1]. Fitting ranges are limited to the bins with  $n_i > 60$ , where  $n_i$  is the number of events in the  $i$ th bin.

Centrality [%]	N	$\hat{k}$	$\hat{n}$	$1/\hat{k}_{dyn}$	$\omega_{dyn}$	$\chi^2/n_d$	p-value [%]	$\chi^2_{LS}/n_d$ with errors:	
						$\chi^2$ ( $n_d$ )		quadrature sum	statistical only
0-5	368271	61.5	19.79	$6.50 \cdot 10^{-3}$	1.13	122.2	0	2.3	23.0
		$\pm 0.6$	$\pm 0.01$	$\pm 0.77 \cdot 10^{-3}$	$\pm 0.02$	4398.3 (36)			
5-10	368869	52.0	16.74	$7.69 \cdot 10^{-3}$	1.13	20.5	0	0.39	3.9
		$\pm 0.5$	$\pm 0.01$	$\pm 0.91 \cdot 10^{-3}$	$\pm 0.02$	613.9 (30)			
10-15	369825	42.3	14.05	$9.46 \cdot 10^{-3}$	1.13	16.2	0	0.43	4.3
		$\pm 0.4$	$\pm 0.01$	$\pm 0.11 \cdot 10^{-2}$	$\pm 0.02$	470.9 (29)			
15-20	369964	35.1	11.77	$1.14 \cdot 10^{-2}$	1.13	11.4	0	0.24	2.4
		$\pm 0.3$	$\pm 0.01$	$\pm 0.13 \cdot 10^{-2}$	$\pm 0.02$	296.8 (26)			
20-25	371752	29.8	9.80	$1.34 \cdot 10^{-2}$	1.13	6.6	0	0.20	2.0
		$\pm 0.3$	$\pm 0.01$	$\pm 0.16 \cdot 10^{-2}$	$\pm 0.02$	370.4 (23)			
25-30	368708	25.6	8.14	$1.56 \cdot 10^{-2}$	1.13	42.7	0	0.21	2.1
		$\pm 0.3$	$\pm 0.01$	$\pm 0.18 \cdot 10^{-2}$	$\pm 0.01$	853.2 (20)			
30-35	367869	22.6	6.72	$1.77 \cdot 10^{-2}$	1.12	126.4	0	0.62	6.2
		$\pm 0.2$	$\pm 0.005$	$\pm 0.21 \cdot 10^{-2}$	$\pm 0.01$	2274.4 (18)			

The results of this analysis show that this is too much. This is connected with the meaning of the formulation of a hypothesis. If the hypothesis is true, it means that the form of a distribution postulated by the hypothesis is exact for the whole population. Thus for the very large samples (as in all PHENIX cases) the measured distribution should be very close to that postulated. The performed analysis has shown that the PHENIX experimental errors are much bigger than the acceptable discrepancies (acceptable for these sample sizes). Therefore  $\chi^2$  and  $\chi^2_{LS}$  test statistics give the opposite answers.

### Acknowledgments

The author thanks Jeffery Mitchell for helpful explanations of the PHENIX data. This work was supported in part by the Polish Ministry of Science and Higher Education under contract No. N N202 231837.

TABLE VII: Results of fitting multiplicity distributions measured by the PHENIX Collaboration in Cu-Cu collisions at  $\sqrt{s_{NN}} = 62.4$  GeV,  $f_{geo} = 0.32 \pm 0.063$  [1]. Fitting ranges are limited to the bins with  $n_i > 5$ , where  $n_i$  is the number of events in the  $i$ th bin.

Centrality [%]	N	$\hat{k}$	$\hat{n}$	$1/\hat{k}_{dyn}$	$\omega_{dyn}$	$\chi^2/n_d$	p-value [%]	$\chi^2_{LS}/n_d$ with errors:	
						$\chi^2$ ( $n_d$ )		quadrature sum	statistical only
0-5	298182	41.6 $\pm 0.4$	13.35 $\pm 0.01$	$7.69 \cdot 10^{-3}$ $\pm 0.15 \cdot 10^{-2}$	1.10 $\pm 0.02$	9.3 279.9 (30)	0	0.65	6.5
5-10	307150	26.5 $\pm 0.2$	11.67 $\pm 0.01$	$1.21 \cdot 10^{-2}$ $\pm 0.24 \cdot 10^{-2}$	1.14 $\pm 0.03$	9.7 290.7 (30)	0	0.78	7.8
10-15	309874	20.5 $\pm 0.2$	9.90 $\pm 0.01$	$1.56 \cdot 10^{-2}$ $\pm 0.31 \cdot 10^{-2}$	1.15 $\pm 0.03$	9.3 261.1 (28)	0	4.4	44.0
15-20	312530	17.8 $\pm 0.1$	8.27 $\pm 0.01$	$1.80 \cdot 10^{-2}$ $\pm 0.36 \cdot 10^{-2}$	1.15 $\pm 0.03$	26.0 677.1 (26)	0	31.6	316.1
20-25	312884	16.0 $\pm 0.1$	6.89 $\pm 0.01$	$1.99 \cdot 10^{-2}$ $\pm 0.39 \cdot 10^{-2}$	1.14 $\pm 0.03$	75.8 1744.0 (23)	0	80.9	809.3

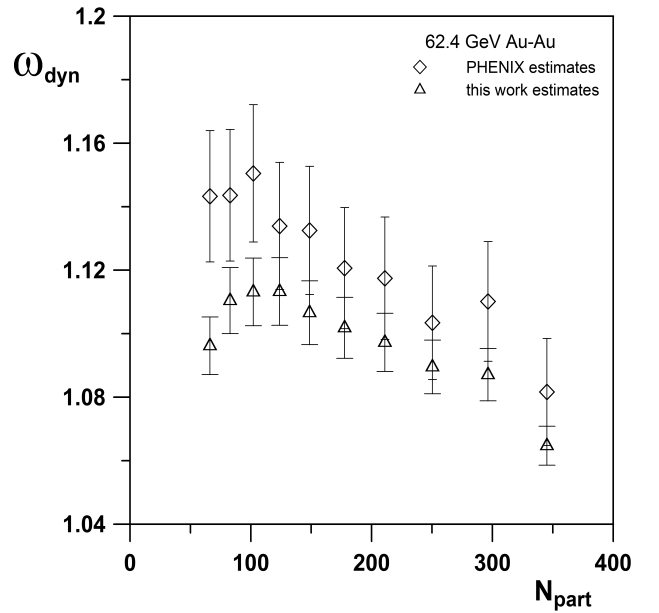
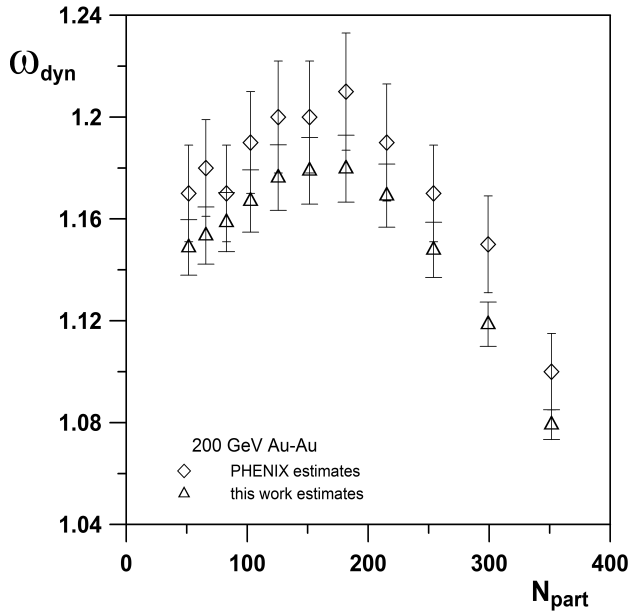


FIG. 7: Scaled variance for 200 GeV (left) and 62.4 GeV (right) Au-Au collisions. PHENIX estimates are from [1]. This work estimates are for the cases with ranges limited to the bins where  $n_i > 60$ , see Tables II and IV.

TABLE VIII: Results of fitting multiplicity distributions measured by the PHENIX Collaboration in Cu-Cu collisions at  $\sqrt{s_{NN}} = 62.4$  GeV,  $f_{geo} = 0.32 \pm 0.063$  [1]. Fitting ranges are limited to the bins with  $n_i > 60$ , where  $n_i$  is the number of events in the  $i$ th bin.

Centrality [%]	N	$\hat{k}$	$\hat{n}$	$1/\hat{k}_{dyn}$	$\omega_{dyn}$	$\chi^2/n_d$	p-value [%]	$\chi^2_{LS}/n_d$ with errors:	
						$\chi^2$ ( $n_d$ )		quadrature sum	statistical only
0-5	298131	42.0 $\pm 0.5$	13.35 $\pm 0.01$	$7.62 \cdot 10^{-3}$ $\pm 0.15 \cdot 10^{-2}$	1.10 $\pm 0.02$	14.7 411.9 (28)	0	0.67	6.7
5-10	307061	26.8 $\pm 0.2$	11.66 $\pm 0.01$	$1.19 \cdot 10^{-2}$ $\pm 0.24 \cdot 10^{-2}$	1.14 $\pm 0.03$	19.7 512.5 (26)	0	0.86	8.6
10-15	309798	20.7 $\pm 0.2$	9.90 $\pm 0.01$	$1.54 \cdot 10^{-2}$ $\pm 0.30 \cdot 10^{-2}$	1.15 $\pm 0.03$	19.4 465.5 (24)	0	0.38	3.8
15-20	312434	18.0 $\pm 0.1$	8.27 $\pm 0.01$	$1.78 \cdot 10^{-2}$ $\pm 0.35 \cdot 10^{-2}$	1.15 $\pm 0.03$	46.5 976.4 (21)	0	0.40	4.0
20-25	312758	16.3 $\pm 0.1$	6.89 $\pm 0.01$	$1.96 \cdot 10^{-2}$ $\pm 0.39 \cdot 10^{-2}$	1.14 $\pm 0.03$	118.1 2243.4 (19)	0	0.63	6.3

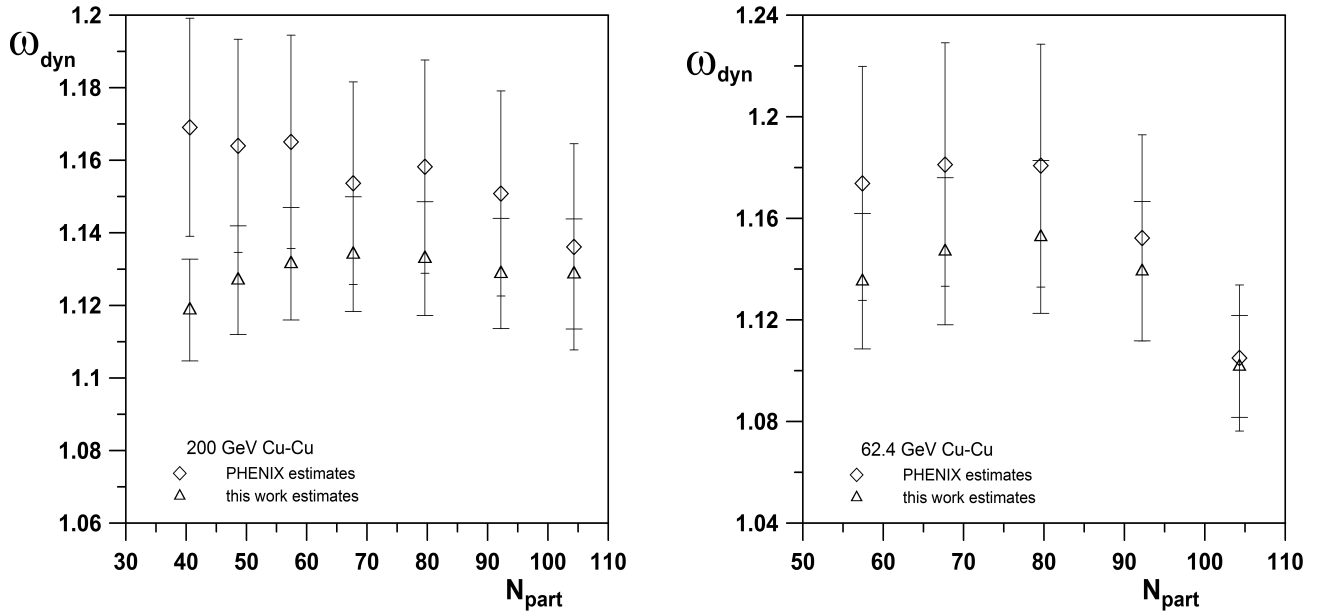


FIG. 8: Scaled variance for 200 GeV (left) and 62.4 GeV (right) Cu-Cu collisions. PHENIX estimates are from [1]. This work estimates are for the cases with ranges limited to the bins where  $n_i > 60$ , see Tables VI and VIII.

### Appendix A

Dropping terms not depending on the parameters in Eq. (17), one obtains the following form for the log-likelihood function under consideration:

$$\ln L(\mathbf{Y} \mid p, k) = N \sum_{i=1}^m P_i^{ex} \ln P(Y_i; p, k) . \quad (\text{A1})$$

Since the logarithm of the NBD is given by

$$\begin{aligned} & \ln P(n; p, k) \\ &= \sum_{j=1}^n \ln(k + j - 1) + n \ln(1 - p) + k \ln p - \ln(n!) , \end{aligned} \quad (\text{A2})$$

the necessary conditions for the existence of the maximum have the following form:

$$\begin{aligned} & \frac{\partial}{\partial p} \ln L(\mathbf{Y} \mid p, k) \\ &= N \sum_{i=1}^m P_i^{ex} \left[ -Y_i \frac{1}{1-p} + \frac{k}{p} \right] \\ &= N \left[ -\frac{1}{1-p} \sum_{i=1}^m P_i^{ex} Y_i + \frac{k}{p} \sum_{i=1}^m P_i^{ex} \right] \\ &= N \left[ -\frac{1}{1-p} \langle N_{ch} \rangle + \frac{k}{p} \right] = 0 , \end{aligned} \quad (\text{A3})$$

$$\begin{aligned} & \frac{\partial}{\partial k} \ln L(\mathbf{Y} \mid p, k) \\ &= N \sum_{i=1}^m P_i^{ex} \left[ \sum_{j=1}^{Y_i} \frac{1}{k + j - 1} + \ln p \right] \\ &= N \left[ \sum_{i=1}^m P_i^{ex} \sum_{j=1}^{Y_i} \frac{1}{k + j - 1} + \ln p \right] = 0 , \end{aligned} \quad (\text{A4})$$

where the sum over  $j$  is 0 if  $Y_i = 0$ .

From Eqs. (A3) and (2) one can obtain:

$$\langle N_{ch} \rangle = \frac{k(1-p)}{p} = \bar{n} . \quad (\text{A5})$$

Expressing  $p$  as a function of  $k$  and  $\langle N_{ch} \rangle$

$$\frac{1}{p} = \frac{\langle N_{ch} \rangle}{k} + 1 , \quad (\text{A6})$$

and substituting it to Eq. (A4) the equation which determines  $\hat{k}$  is obtained:

$$\frac{\partial}{\partial k} \ln L(\mathbf{Y} \mid p, k)$$



$$= N \left[ \sum_{i=1}^m P_i^{ex} \sum_{j=1}^{Y_i} \frac{1}{k+j-1} - \ln \left( 1 + \frac{\langle N_{ch} \rangle}{k} \right) \right] = 0. \quad (\text{A7})$$

The above equation can be solved numerically. Having obtained  $\hat{k}$  and substituting it into Eq. (A6)  $\hat{p}$  is derived.

## Appendix B: Wilks's theorem

Let  $X$  be a random variable with p.d.f  $f(X, \theta)$ , which depends on parameters  $\theta = \{\theta_1, \theta_2, \dots, \theta_d\} \in \Theta$ , where a parameter space  $\Theta$  is an open set in  $\mathbb{R}^d$ . For the set of  $N$  independent observations of  $X$ ,  $\mathbf{X} = \{X_1, X_2, \dots, X_N\}$ , one can define the likelihood function

$$L(\mathbf{X} | \theta) = \prod_{j=1}^N f(X_j; \theta). \quad (\text{B1})$$

Now consider  $H_0$ , a  $k$ -dimensional subset of  $\Theta$ ,  $k < d$ . Then the maximum likelihood ratio can be defined as

$$\lambda = \frac{\max_{\theta \in H_0} L(\mathbf{X} | \theta)}{\max_{\theta \in \Theta} L(\mathbf{X} | \theta)}. \quad (\text{B2})$$

This is a statistic because it does not depend on parameters  $\theta$  no more, in the numerator and the denominator there are likelihood function values at the ML estimators of parameters  $\theta$  with respect to sets  $H_0$  and  $\Theta$ , respectively.

The Wilks's theorem says that under certain regularity conditions if the hypothesis  $H_0$  is true (*i.e.* it is true that  $\theta \in H_0$ ), then the distribution of the statistic  $-2 \ln \lambda$  converges to a  $\chi^2$  distribution with  $d - k$  degrees of freedom as  $N \rightarrow \infty$  [4, 6]. The proof can be found in Ref. [9]. Note that  $k = 0$  is possible, so one point in the parameter space (one value of the parameter) can be tested as well.

- 
- [1] A. Adare *et al.* [PHENIX Collaboration], *Phys. Rev. C* **78**, 044902 (2008).
  - [2] UA5 Collab. (G. J. Alner *et al.*), *Phys. Lett. B* **160**, 193 (1985).
  - [3] G. Cowan, *Statistical data analysis*, (Oxford University Press, Oxford, 1998)
  - [4] F. James, *Statistical methods in experimental physics*, (World Scientific, Singapore, 2006)
  - [5] S. Baker and R. D. Cousins, *Nucl. Instrum. Meth.* **221**, 437 (1984).
  - [6] P. G. Hoel, *Introduction to mathematical statistics*, 4th edn. (Wiley, New York, 1971)
  - [7] E-802 Collab. (T. Abbott *et al.*), *Phys. Rev. C* **52**, 2663 (1995).
  - [8] J. T. Mitchell, private communication.
  - [9] R. M. Dudley, *18.466 Mathematical Statistics, Spring 2003*, (Massachusetts Institute of Technology: MIT OpenCourseWare), <http://ocw.mit.edu/courses/mathematics/18-466-mathematical-statistics-spring-2003/lecture-notes/>
  - [10] Here, these quantities are distinguished from the experimentally measured the average charged particle multiplicity  $\langle N_{ch} \rangle$  and the variance  $\sigma^2$ .
  - [11] Precisely, because of the energy conservation the number of produced charged particles is limited but the number of collisions is not.
  - [12] Now  $Y_i \neq Y_j$  for  $i \neq j$  and  $i, j = 1, 2, \dots, m$ .